FINAL REPORT

Contract NAS5-98046

Period covered: 31 january 1998 -30 September 2001

Prepared for
NASA/Goddard Space Flight Center

JB Kumer
Principal Investigator

Lockheed Martin Advanced technical center (LMATC)
LOCKHEED MARTIN MISSILES & SPACE COMPANY
Palo Alto, California, 94304-1191

Final Report on NAS5-98046

CLAES Product Improvement by use of GSFC Data Assimilation System

1.0 Motivation: Recent development in chemistry transport models (CTM) and in data assimilation systems (DAS) indicate impressive predictive capability for the movement of airparcels and the chemistry that goes on within these. This project was aimed at exploring the use of this capability to achieve improved retrieval of geophysical parameters from remote sensing data. The specific goal was to improve retrieval of the CLAES CH_4 data obtained during the active north high latitude dynamics event of 18 to 25 February 1992. The model capabilities would be used

- rather than climatology to improve on the first guess and the a-priori fields
- to provide horizontal gradients to include in the retrieval forward model

The retrieval would be implemented with the first forward DAS prediction. The results would feed back to the DAS and a second DAS prediction for first guess, a-priori and gradients would feed to the retrieval. The process would repeat to convergence and then proceed to the next day.

2.0 Approach: This initial work was focussed on 20th February as it was one of the most active in the period, and several anomalous retrieved CH_4 profiles were identified on that day and tagged for special attention. To test improvement that might be realized by the approach above this initial work used the available CTM output for that day to

- provide first guess and the a-priori fields
- horizontal gradients to include in the retrieval forward model

With this input the retrieval was implemented to

- first to retrieve profiles for the specific EMAFs of the anomalous profiles
- second to map the entire day

And the results were examined in order to assess the extent of improvement. The results are discussed below in the **Findings** section 4.0.

3.0 Implementation: The implementation cited above required considerable effort that is described here. First for every EMAF, and each of the blocker filter regions, it is required to compute location data along the line of sight from the CLAES through the atmosphere that could be utilized by the CLAES retrieval software (RSW) to include horizontal gradients into the forward model. The forward model was provided by GATS. The forward model indexing system & the definitions & conventions that are required to correctly apply the model for horizontal temperature gradients are shown in Appendix A, and for mixing ratios in Appendix B. For CLAES the model was applied for 25 lines of sight (LOS). CTM field data at locations on these LOS as described below are required to run the forward model with gradients included. The location data are

- a set of pressure surfaces PRESSURES(IZMAX), the value IZMAX=25 is used by our RSW
- a set of latitudes LAT_n(IZMAX,IZMAX) and longitudes $LONG_n$(IZMAX,IZMAX) along the near side of the rays
- a set of latitudes LAT_f(IZMAX,IZMAX) and longitudes $LONG_f$(IZMAX,IZMAX) along the far side of the rays

In addition there will be a time (in UT) associated with each profile.

Coding needed to be developed to compute these locations. Input for the computation is provided in the instrument data stream by the spacecraft and includes coordinates (R,θ,ϕ) of the tangent point TP, ϕ_{TP} the angle east from north at the LOS from the spacecraft to the TP, and R_{SC} the distance from earth center to the spacecraft. The transforms and procedures that required coding in order to compute the location data are shown in Appendix C.

Next it was required to interpolate the CTM data to these locations, and input these to the CLAES RSW in order to implement the retrieval runs. Two approaches for this were utilized, as it was found the first approach was too cumbersome to be practical.

The first approach involved computation of the location data in files of one record per EMAF. These were emailed to GSFC CTM personnel who developed a routine to interpolate CTM fields to the locations, and store these in a file at a record per EMAF, and email them back to LMATC.

After some experience with this approach we realized that this method for data interface (as described above) with the CTM was cumbersome and not conducive to efficiently handling whole day size data sets. Therefore we requested the GSFC CTM personnel that we switch to a method in which they provide us with a whole day data set and an interpolating FORTRAN reader subroutine that looks like:

```
XRATIO=READ_DAS_OR_CTM(DATE,TIMEINGMT,LAT,LON,PRESSURE,
SPECIEIDENTIFIER).
```

The GSFC personnel did this, and we worked this into our code at expense of more effort than we had expected because the GSFC data and code were UNIX based as compared to our VMS. But once working it was invaluable in development and debugging as it facilitated many short runs in any part of the data day, that same working day, rather than waiting for turn around with the GSFC personnel. Another factor in the improvement is, since both the LMATC and GSFC personnel had other commitments, and work in different time zones, turn around by the old method was at times lengthy.

Also, in the process of working the second method into the RSW we found that we had an error in the way we used the first method that produced a one EMAF offset in the data returned from the CTM. This lead to incorrect impression as described in our progress report #4 that the mixing ratio gradients had a significant effect (relative to the effect of temperature gradients) on the retrieval. This misimpression was corrected in our report #5 and verified in our report #6.

Further coding changes were necessary to correctly use the gradients in the CLAES RSW. The starting RSW version would use one call to compute radiance on all ray paths. This in two different program modules, and with some differences within these modules, for the temperature retrieval and the mixing ratio retrievals, respectively. As described in Appendices A and B the gradient input needs to be calculate for each LOS, and to implement this required changing the forward model over to calculate one LOS at a time in a loop, and at the same time to keep track and account for the differences in the two modules.

4.0 Findings: Here we summarize the findings that have been discussed in more detail in our progress reports. First, use of the CTM for a-priori and first guess did not perceptibly change retrieval results.

Next, including the gradients in the forward model did induce noticeable effects. In case of the anomalous CH_4 profiles, it was found that these were noticeably changed by using the CTM gradients. The major change was due to the temperature gradients. Adding in the CH_4 gradients had a perceptible but considerably smaller effect. This is illustrated by figure 3 in our progress report #6. Figure 3, and other relevant figures from report #6 are included here in Appendix D for the readers convenience. Qualitatively most noticeable is the increase in the CH_4 at the lowest altitude. However, this was not consistent from profile to profile, as shown by the zonal mean comparisons that are discussed below.

Zonal means of baseline retrieved CH_4 and of CH_4 retrieved by use of forward model that includes CTM temperature and CH_4 horizontal gradients are compared on figures 5 and 6 of (report #6, ie., Appendix D). These are labeled CTMTANDXGRADS and BASELINE, respectively. Two obvious things we learned from these were

1 there is little difference in a zonal mean sense in the retrieved product

2 The CTMTANDXGRADS is spiky at the high altitudes. This probably can be traced to a few EMAFS. There may also be some 'spike' problems at the lower altitudes.

Next on figures 7 and 8 of (report #6, ie., Appendix D) is shown a comparison of CH_4 retrieval CTMTANDXGRADS and CTMONLYGRADS, where in the latter the CH_4 gradients are not included. Again, it can be seen that the species gradients have only a small effect. And as also discussed in our progress report #6, these effects were noted in N_2O retrievals too.

Surface maps on 6.8 and 3.1 mb were also made and these reveal interesting features in comparison between the various CLAES retrievals (CTMTANDXGRADS and BASELINE) and the CTM fields (also mapped). These are shown on figures 11 through 30 in our progress report #6 and are discussed thoroughly there. The major points made from the CH_4 6.8 mb maps were

1 The various CLAES retrievals are notably more similar to one another than to the mapped CTM.

2 The CLAES retrievals show much more ascending to descending difference than does the CTM

3 the CLAES products show an enhanced band going from zero longitude & 50N, to about 100 longitude and 70N that is not present in the CTM, the CLAES north projecting enhancement at about 260 longitude is displaced towards about 220 longitude in the CTM, and more enhanced. There is a small common northward projecting enhancement common to CLAES and CTM at ~ 130 longitude. In CLAES this latter seems more to be part of the major enhancement emanating from zero longitude towards the northeast.

These points are generally supported by maps of N_2O on the 6.8 mb surface. The CH_4 3.1 mb maps add the following

1 generally the CTM has smaller values now than the CLAES products

2 the CTM & CLAES morphology look more alike at 3.1 mb than at 6.8 mb

3 the relative enhancement of the feature protruding to the northeast from about zero long & 40N(call it F1), and of the feature protruding almost due north at about 100 long (call it F2) has reversed. The feature centered at about 200 long and 70N (call it F3) now clearly appears to be an extension of the feature F2 that goes on across the polar region.

4 The CTMTANDXGRADS has noticeably less ascending to descending difference than does the BASELINE

N_2O maps @ 3.1 mb are consistent with the conclusions drawn with regard to CH_4 as discussed above.

5.0 Progress since time of report #6: Since then we have focussed on trying to gain some understanding of the ascending vs descending differences in mapped CH_4.

First we looked at the mapped temperature to see if similar ascending vs descending difference is occurring, and found that it is not as shown by figures 1 and 2 below.

Next we checked our V9 product to see if it had similar ascending vs descending differences in mapped CH_4. The V9 CH_4 mapped to the 6.8 mb surface is shown on figures 3 and 4 below. It is seen to look notably more like the CTMTANDXGRADS on figures 13 and 14 of (report #6, ie., Appendix D) than the BASELINE on figures 11 and 12 of (report #6, ie., Appendix D). The BASELINE and the V9 differ mostly in that the BASELINE uses temperature gradients derived from NMC for use in the species forward model, whereas the V9 uses a GATS routine to derive these from the retrieved CLAES temperature. Prior to this study no attempt had been made to determine which approach is best, as the differences in retrieved product are typically small compared to other effects. The BASELINE approach was used in development work involved in this study because it is easier to implement, and more conducive for development applications that involve

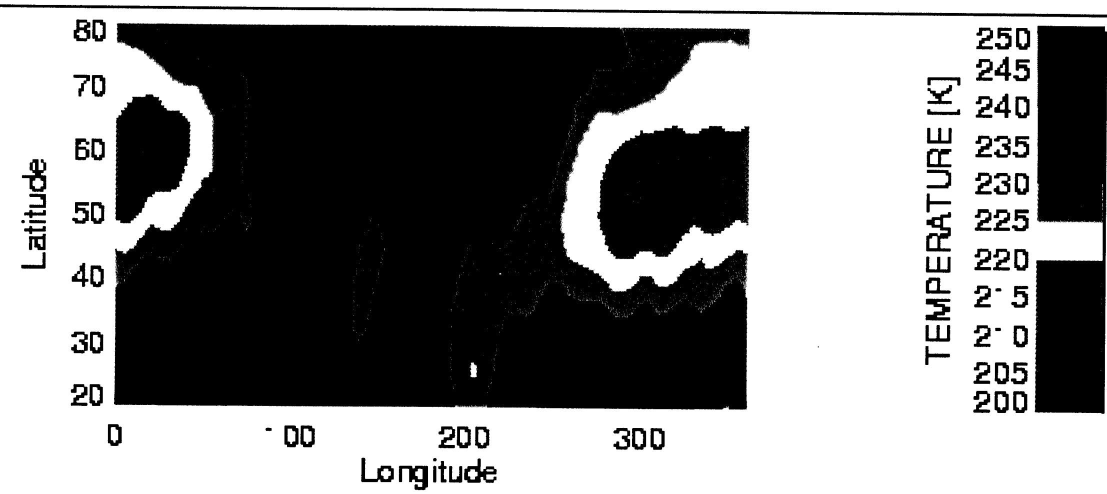


Figure 1 temperature mapped on to 6.8 mb for the ascending CORRECTED BASELINE case

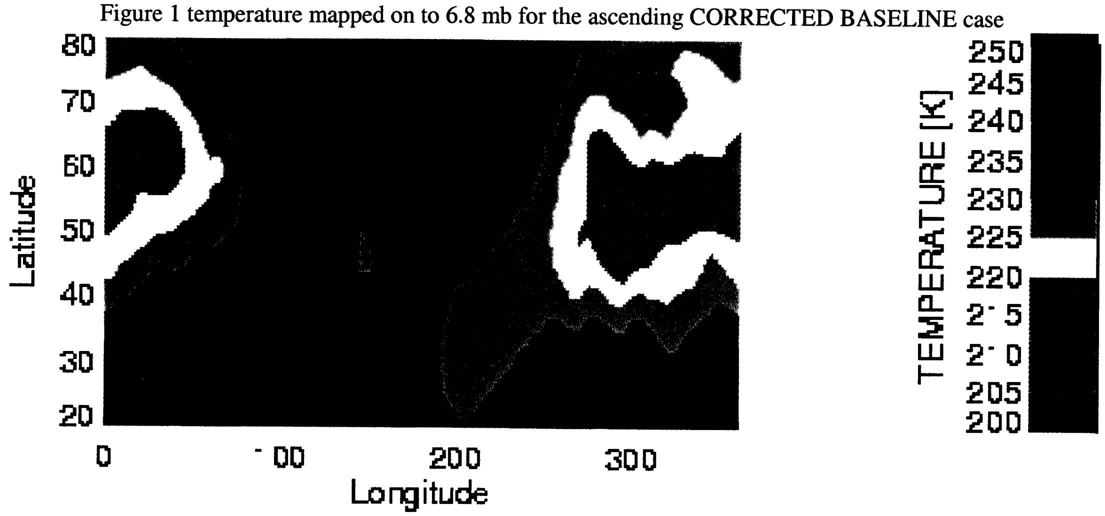


Figure 2 temperature mapped on to 6.8 mb for the descending CORRECTED BASELINE case many short runs of only a few EMAFs, which is the mode this effort necessarily began in. However, the odd man out character of the BASELINE motivated us to look for an error in the way we were deriving temperature gradients from the NMC fields. After considerable painful effort in going through the code for that, and making many test runs to make sure it was the BASELINE, and not the CTMTANDXGRADS, that was in err, we did indeed find a bug in the BASELINE implementation of deriving error from the NMC fields. When fixed we get the CORRECTED_BASELINE result as shown on figures 5 and 6 here. These now look more similar to the CTMTANDXGRADS and the V9, although it's clear the BASELINE and V9 look more like each other than they do the CTMTANDXGRADS. It's also seen that the ascending vs descending differences are the least in the CTMTANDXGRADS, but are still quite noticeable.

In summary, our work since the time of report #6 has focussed on understanding why there are ascending vs descending differences in the mapped CH_4 at 6.8 mb. It has shown that this is not present in the retrieved temperature, so that is not a factor. An

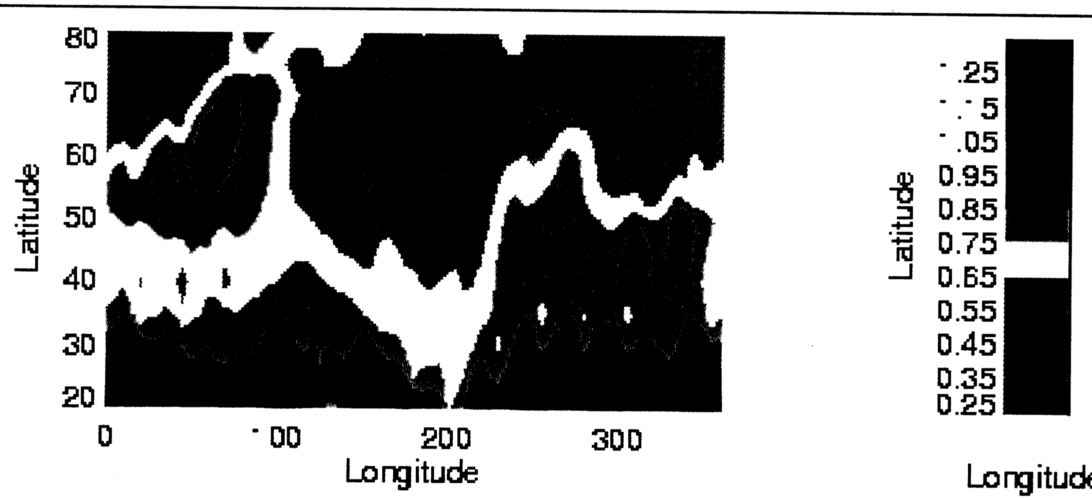


Figure 3 CH_4 mapped on to 6.8 mb for the ascending V9 case

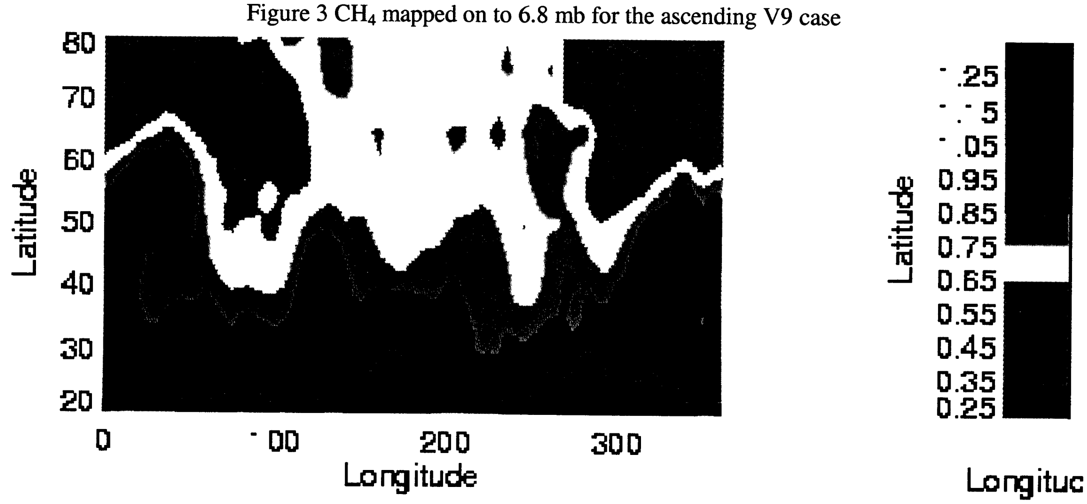


Figure 4 CH_4 mapped on to 6.8 mb for the descending V9 case unanticipated result was a validation of the temperature gradients that are used in the V9 species forward model.

6.0 Top level result and potential follow on work: The top level result of this study was that use of the CTM gradients in the forward model does not dramatically change the CLAES CH_4 product. This suggests using the DAS together with the retrieval software in an iterative scheme to implement CLAES CH_4 retrieval also would probably not make a dramatic change.

This is not to say that this approach is not the way to go in future remote sounding application, as it is to say that its utility is limited for the specific case of CLAES CH_4. The major reason for this is that retrieval should be most sensitive to horizontal gradients at altitudes well below the maxima in species such as O_3, HNO_3, $ClONO_2$, etc. This is not the case for CH_4 and is most likely the reason no dramatic effect was seen to occur by use of the CTM gradients.

Follow on work would address two areas. First wrap up the work on CH_4 including

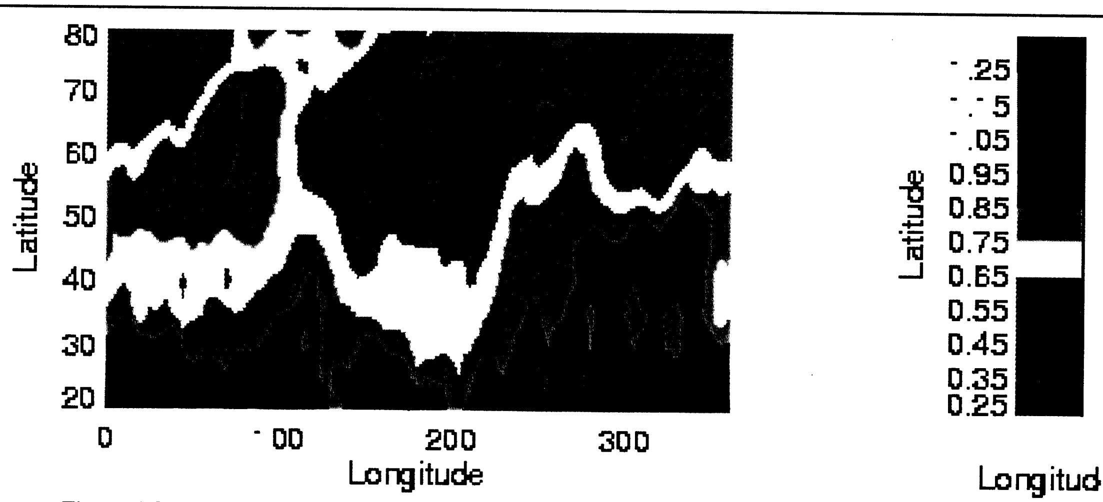


Figure 5 CH_4 mapped on to 6.8 mb for the ascending CORRECTED BASELINE case

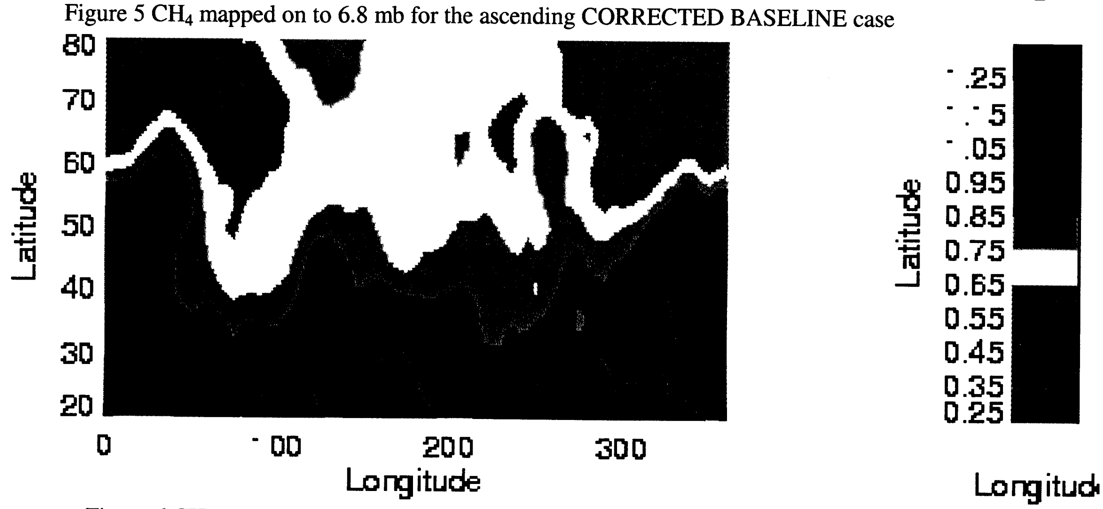


Figure 6 CH_4 mapped on to 6.8 mb for the descending CORRECTED BASELINE case tracking down the reasons for the spikines that is seen in the zonal means at high altitude. Second, continue to address the problem of understanding and correcting for the non physical ascending vs descending differences in the CH_4. Regarding the latter, there are portions of the CLAES retrieved H_2O data set that provide an even more striking case, and it might be easier to solve the problem in working with these data, rather than the CH_4. Third, test the gradients effects for the species that are considered more likely to be affected such as O_3, HNO_3, $ClONO_2$, etc.

Appendix A: Location Data and temperature gradients definition required for inclusion of horizontal gradients into forward model

Based on email discussion with Tom marshall this is the correct approach.

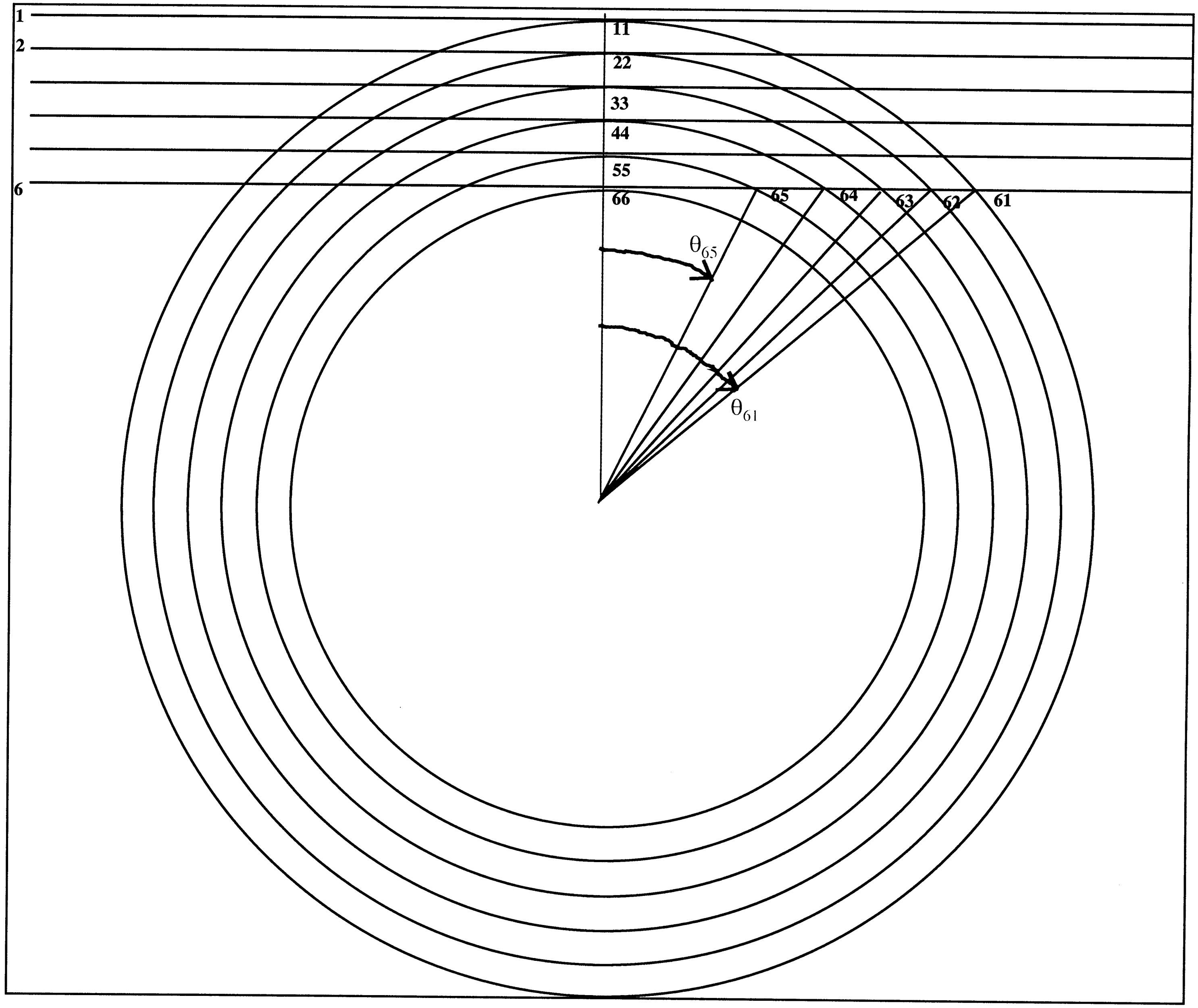
He verified that GR(1,N) refers to the side towards the spacecraft and that GR(2,N) refers to the other side from the spacecraft.

The available input data are.

- The temperatures T_{66}, T_{55},... T_{11} at the points along the line from earth center through the tangent points
- The temperature T_{66}, T_{65},... T_{61} at the points along the LOS to the space craft
- The angles θ_{65} from the tangent point to the intersection 65, θ_{64} from the tangent point to the intersection 64,... θ_{61} from the tangent point to the intersection 61,

Based on our email exchange it is correct usage to compute GR(1,N) for the 6th ray as below..

$GR(1,1)=(T_{61}-T_{11})/\theta_{61}$
$GR(1,2)=(T_{62}-T_{22})/\theta_{62}$
$GR(1,3)=(T_{63}-T_{33})/\theta_{63}$
$GR(1,4)=(T_{64}-T_{44})/\theta_{64}$
$GR(1,5)=(T_{65}-T_{55})/\theta_{65}$
$GR(1,N)=0$ for $N > 5$



Appendix B: Location Data and mixing ratio gradients definition required for inclusion of horizontal gradients into forward model

Based on email discussion with Tom Marshall this may be the correct approach for mixing ratio.

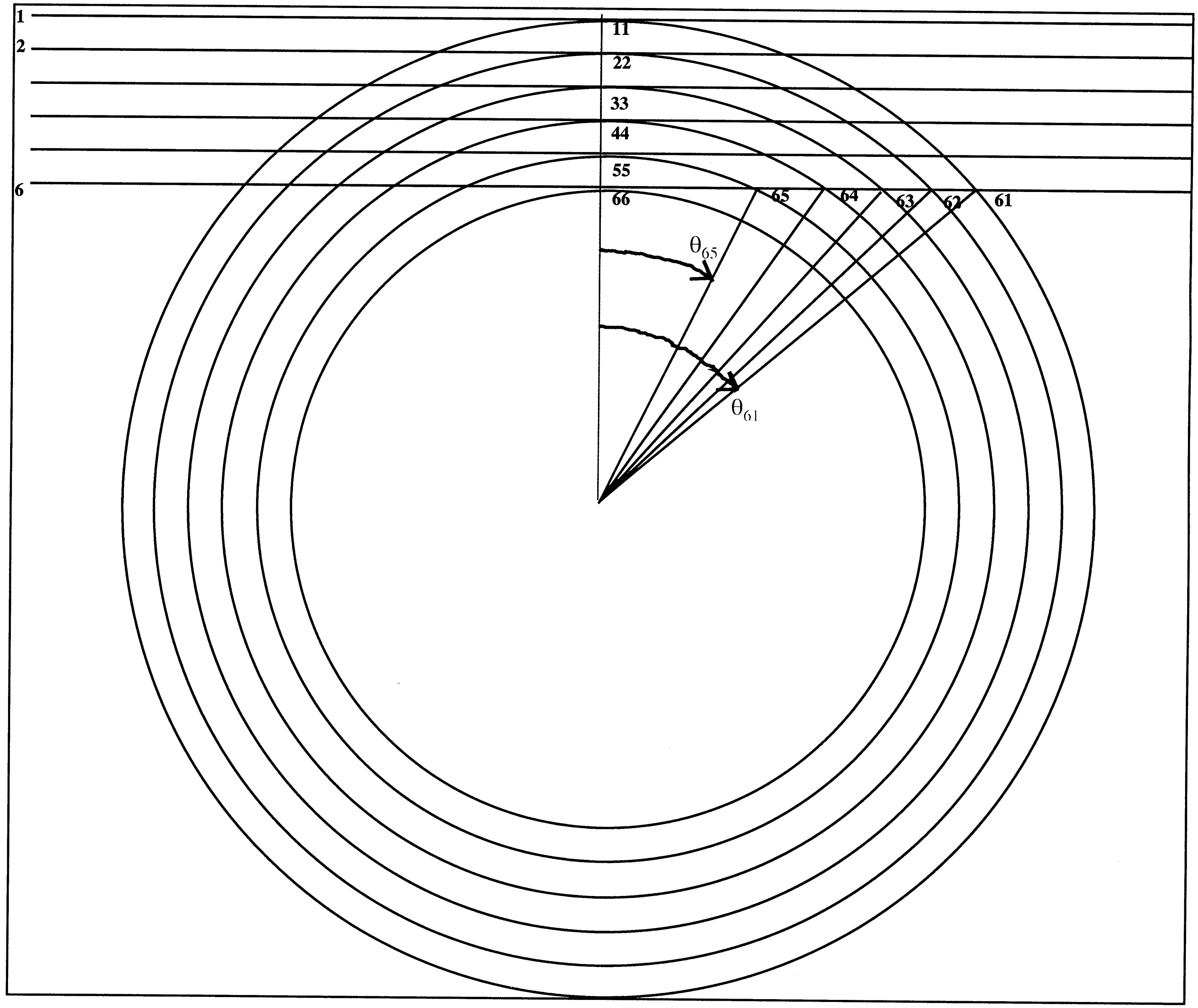
He verified that QG(1,N) refers to the side towards the spacecraft and that QG(2,N) refers to the other side from the spacecraft.

The available input data are.

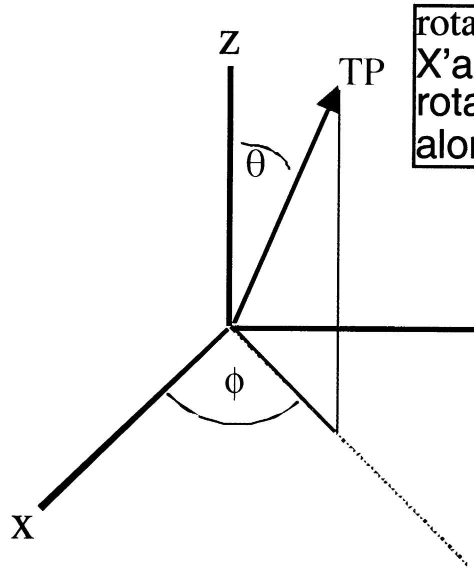
- The mixing ratios X_{66}, X_{55},... X_{11} at the points along the line from earth center through the tangent points
- The mixing ratio X_{66}, X_{65},... X_{61} at the points along the LOS to the space craft
- The angles θ_{65} from the tangent point to the intersection 65, θ_{64} from the tangent point to the intersection 64,... θ_{61} from the tangent point to the intersection 61,

Based on our email exchange it is correct usage to compute QG(1,N) for the 6th ray as below..

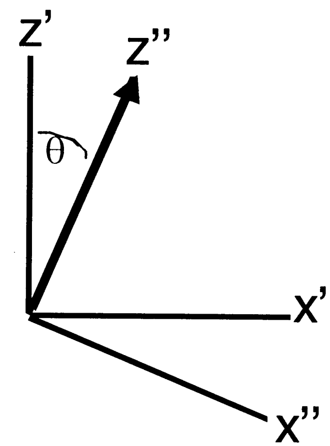
$QG(1,1)=(X_{61}/X_{11}-1)/\theta_{61}$
$QG(1,2)=(X_{62}/X_{22}-1)/\theta_{62}$
$QG(1,3)=(X_{63}/X_{33}-1)/\theta_{63}$
$QG(1,4)=(X_{64}/X_{44}-1)/\theta_{64}$
$QG(1,5)=(X_{65}/X_{55}-1)/\theta_{65}$
$QG(1,N)=0$ for $N > 5$

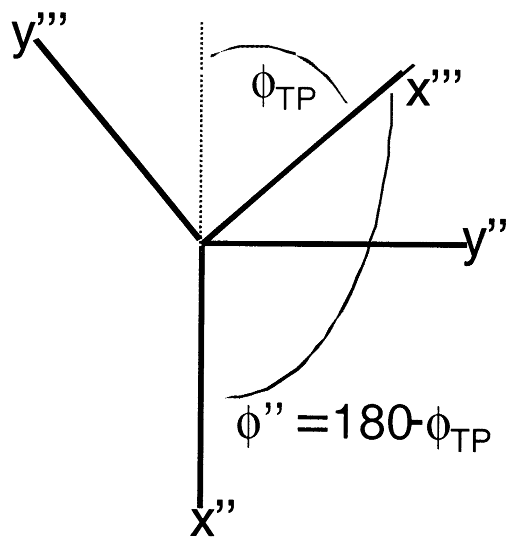


Appendix C: Calculating the locations for CTM field variables along CLAES line of sight (LOS)



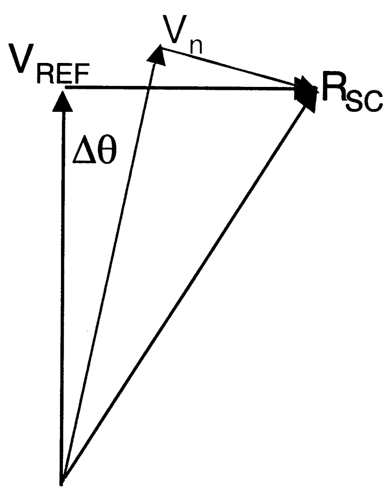
rotate about z by longitude ϕ to form the prime coordinate system, with X'axis along the tangent point (TP) projection to the XY plane and th rotate about Y' by latitude θ to form the system which has the Z'' axis along line from the earth center to the tangent point TP as shown be



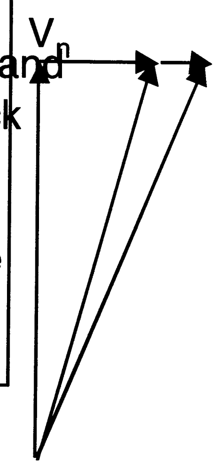


Now, looking straight down along the Z'' axis is shown to th left. The ϕ_{TP} is the angle east from north of the LOS to the s craft (SC). Rotating by ϕ'' about Z'' as shown points the X''' axis at the SC

The geometry shown on the previous slide applies for the reference TP at the point exactly between the 10th & 11th real detectors. The vector from earth center to this reference TP we call V_{REF}. We actually want the point directly in the middle of the layer for the 1st through 25th forward model vertical layers. These go from top to bottom and define virtual detectors. Counting down from the top the 15th & 16th virtual detectors correspond to the 11th & 10th real detectors (which are counted **up** from the bottom). The vector from earth center that places the TP in the center of the nth virtual detector we call V_n. It differs in length from the V_{REF} by ~ $z_n-(z_{10}+z_{11})/2$ where z_n is the tangent height of the nth virtual detector.For n < 16 a positive rotation about YÕÕÕ is required to point V_n correctly. The rotation is through a small angle $\Delta\theta$ given by $\Delta\theta = \mathrm{ARCSIN}(v_n/R_{SC}) - \mathrm{ARCSIN}(v_{REF}/R_{SC})$ where R_{SC} is the distance to earth center from the S/C, v_n is the distance to the center of the nth virtual detector and v_{REF} is the distance to the reference TP. The correct sign (+ or -) for the angle $\Delta\theta$ is guaranteed by the definition of the v_n. See the picture to the left below. The series of rotations that go from the (θ,ϕ) coordinates to the coordinates of the nth virtual detector are $R_{y'}(\Delta\theta)R_{z''}(\phi'')R_{y'}(\theta)R_z(\phi)$. The transformation of the vector to the nth virtual detector back to standard coordinates is given by $R_z(-\phi)R_{y'}(-\theta)R_{z''}(-\phi'')$.



Next, we need define steps along the vector from V_n to R_{SC} that intersect the pressure levels m < n. These are spaced by $x_{nm}=(2v_n(z_m - z_n))^{1/2}$ and are illustrated to the right for m=n-1 and n-2. The vector with components $(x_{nm},0,v_n)$ is transformed back to the vector (x,y,z) in the $\phi(\theta)$ system, and the θ_{nm} = ARCCOS(z/r) where $r=(x^2+y^2+z^2)^{1/2}$, and ϕ_{nm} solves $x=\rho\mathrm{COS}(\phi_{nm})$ and $y=\rho\mathrm{SIN}(\phi_{nm})$, and where $\rho=(x^2+y^2)^{1/2}$. These are the near side location data named LAT_n and $LONG_n$. Transform $(-x_{nm},0,v_n)$ to obtain the far side location data.



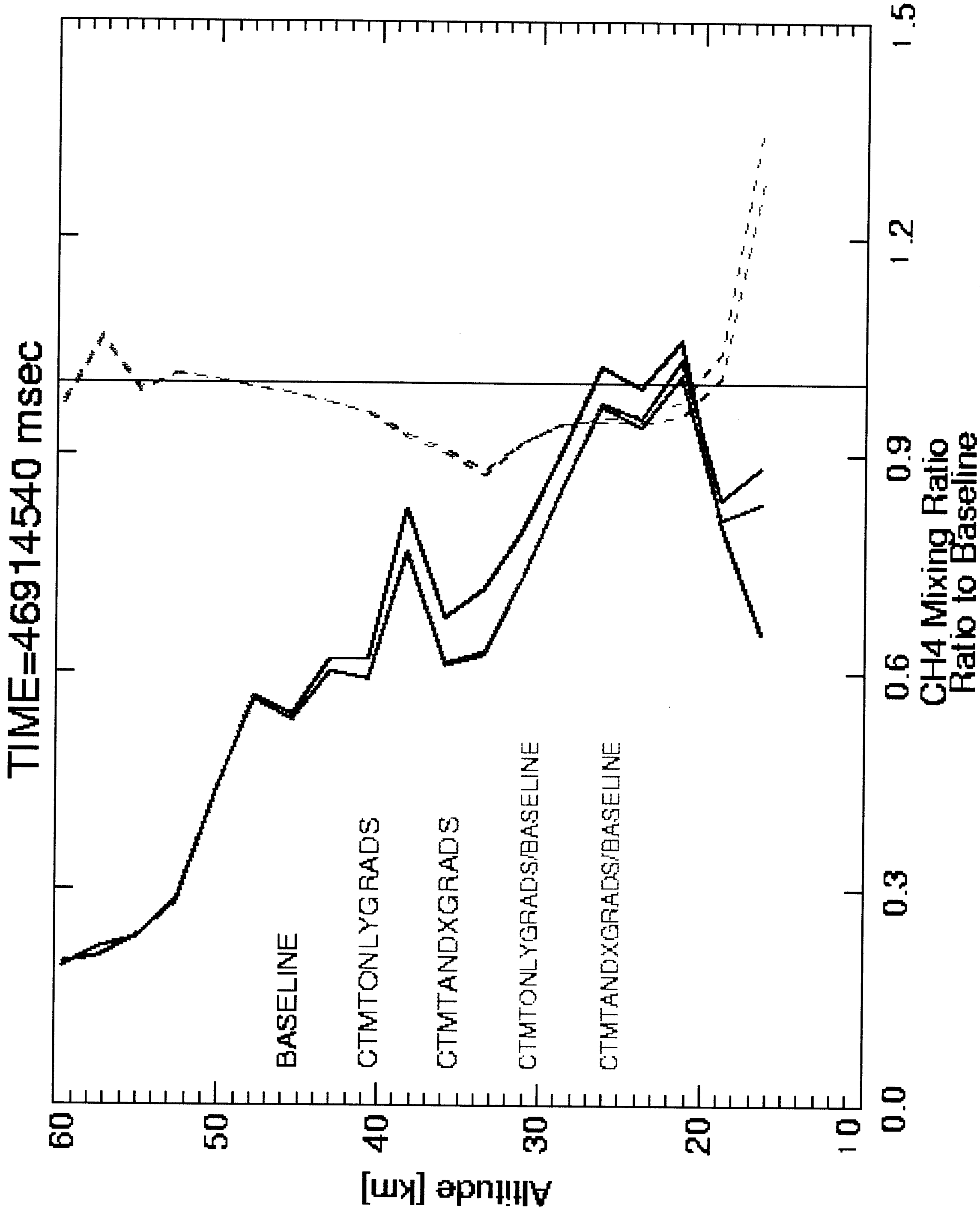
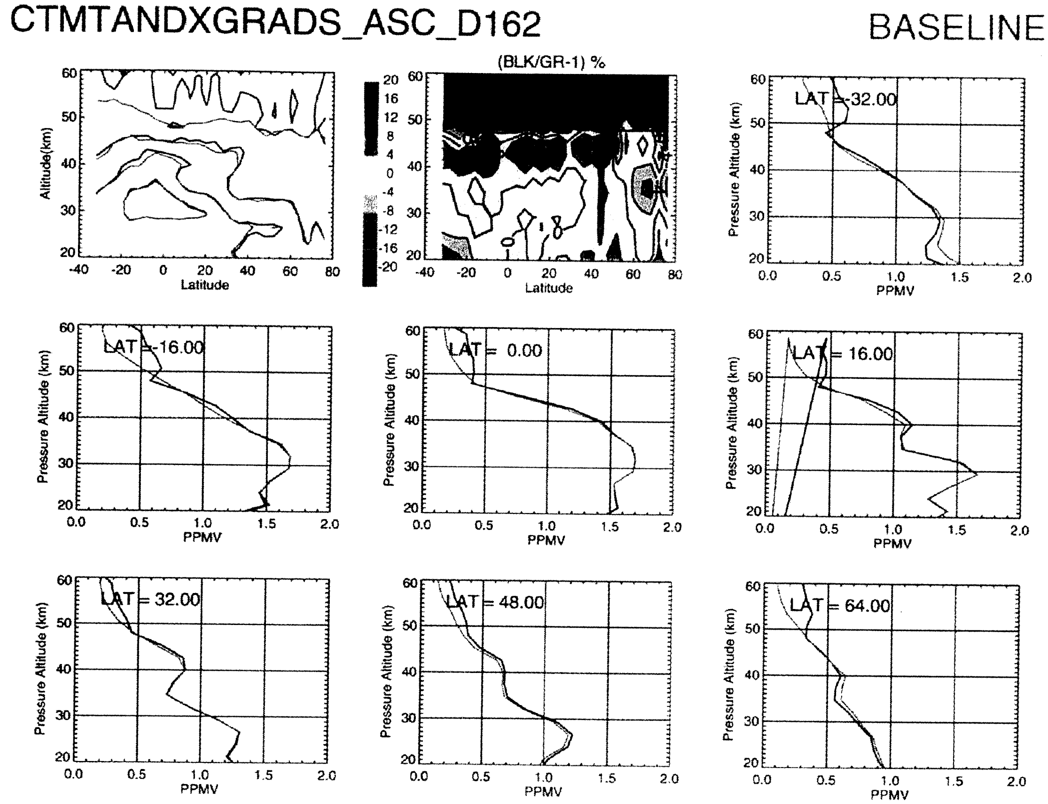
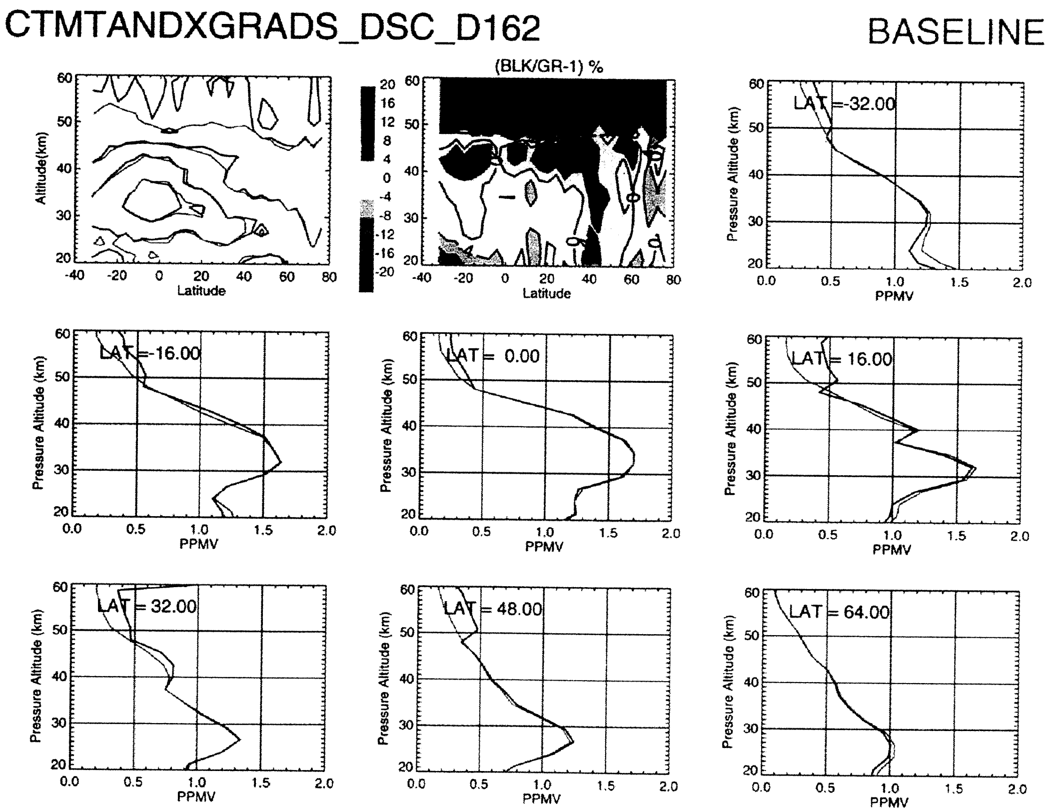
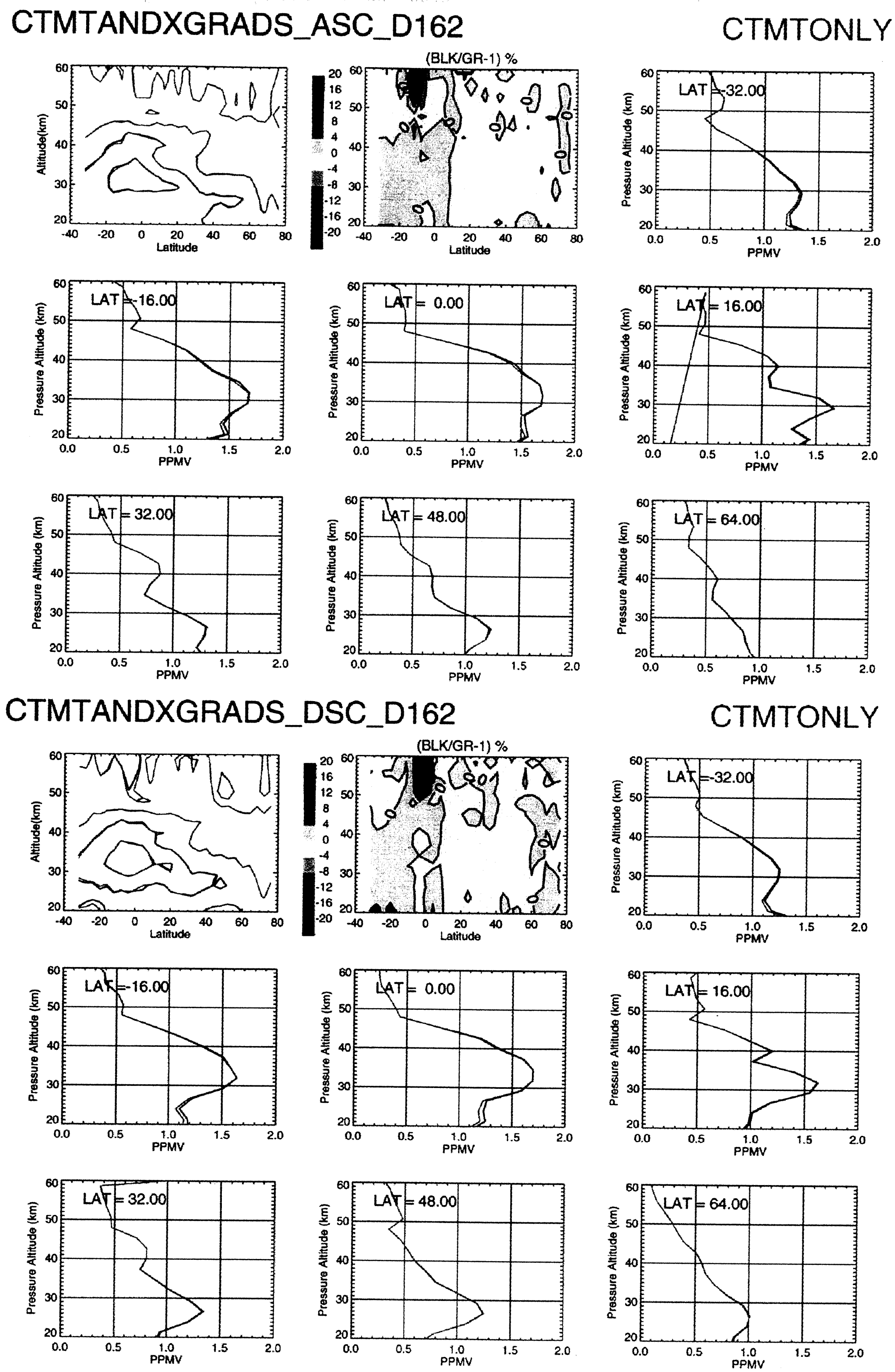


Figure 3 CH_4 retrieval comparisons for EMAF 46914540

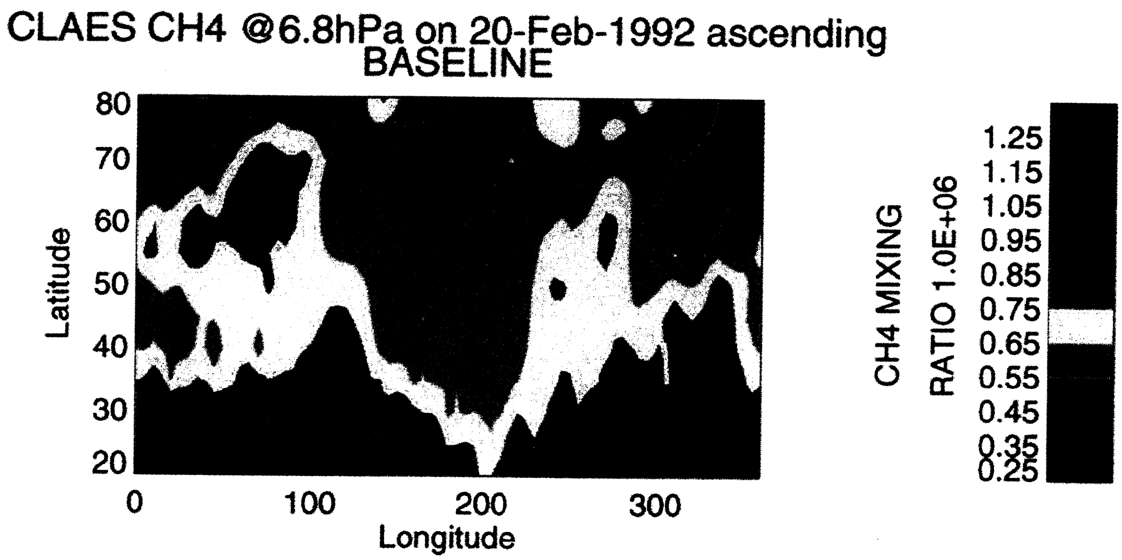


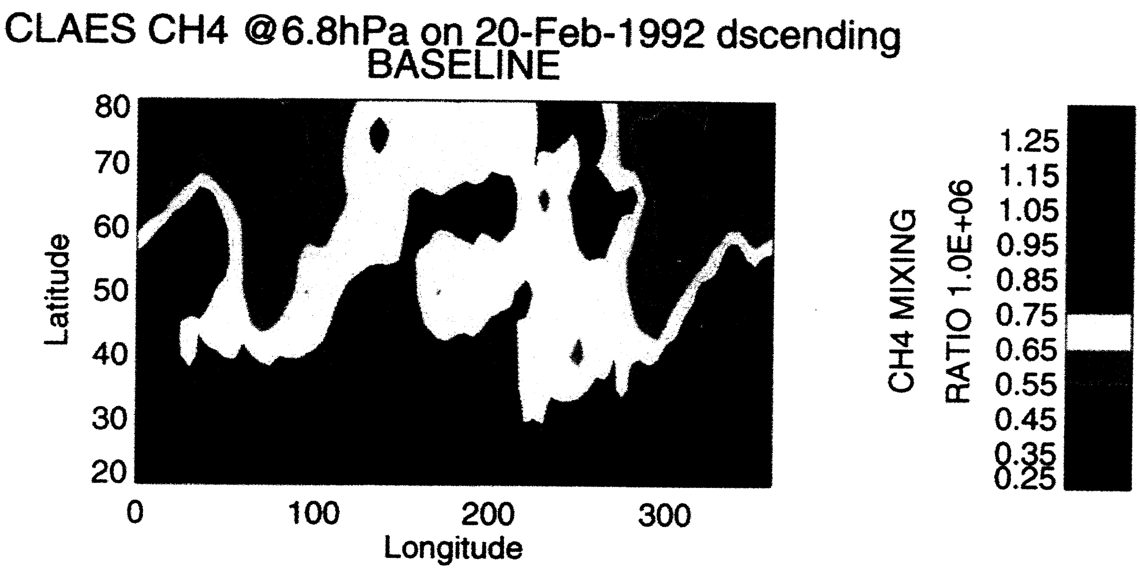


Figures 5 and 6 CH_4 zonal mean comparison CTMTANDXGRADS and BASELINE

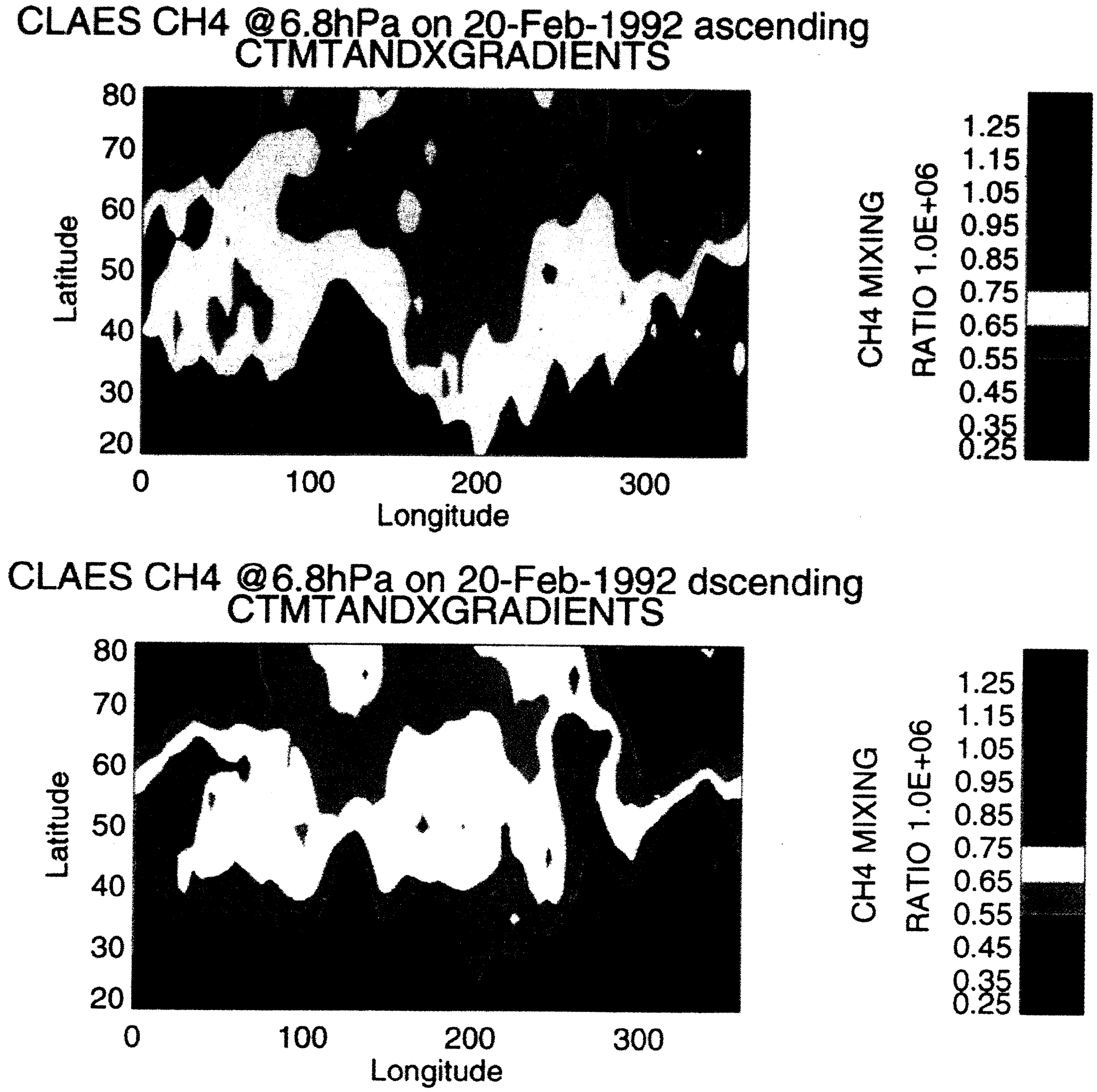


Figures 7 and 8 CH_4 zonal mean comparison CTMTANDXGRADS and CTMTONLY

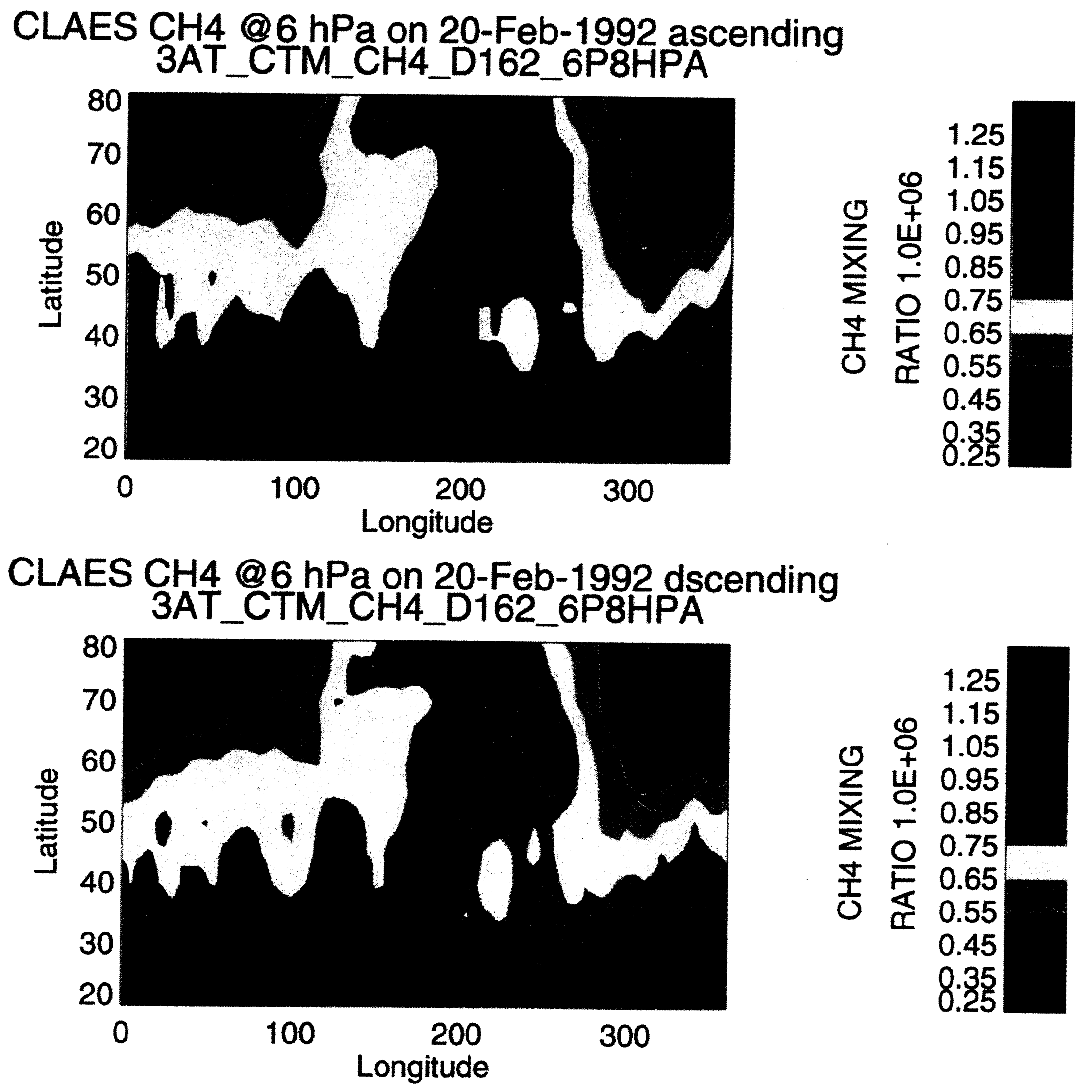




Figures 11 and 12 CH_4 baseline on 6.8 mb



Figures 13 and 14 CH_4 CTMTANDXGRADS on 6.8 mb



Figures 15 and 16 CH_4 CTM on 6.8 mb